Estimating the number of successful spawners in a steelhead population from a sample of offspring

Jesse McCane, Pacific States Marine Fisheries Commission, Eagle, ID 83616

Craig A. Steele, Pacific States Marine Fisheries Commission, Eagle, ID 83616

Christine Kozfkay, Idaho Department of Fish and Game, Eagle, ID 83616

Matthew Campbell, Idaho Department of Fish and Game, Eagle, ID 83616

Background and Study Rationale

Determining genealogical relationships among individuals within a sample of fish without parental information has many important conservation and management applications. Although several methods are currently available to estimate these relationships, they are computationally intensive and require highly informative genetic markers that are free of genotyping errors and mutations in order to provide accurate reconstruction. In this study, we used a powerful new set of 188 single nucleotide polymorphic markers (SNPs) developed for steelhead to compare the effectiveness of existing sibling reconstruction computer programs in correctly reconstructing family relationships using SNP genotypes from hatchery offspring with known parentage. We also compared estimates of effective population size (N_e), a key parameter in assessing both the short- and long-term viability of populations, derived from sibship assignment and linkage disequilibrium against the known N_e values calculated from each parent/offspring group. Verification of the accuracy and performance of these methods could provide opportunities for assessing patterns of dispersal and estimating N_e in future research programs where field collections only target one or two cohorts and may not include parents.

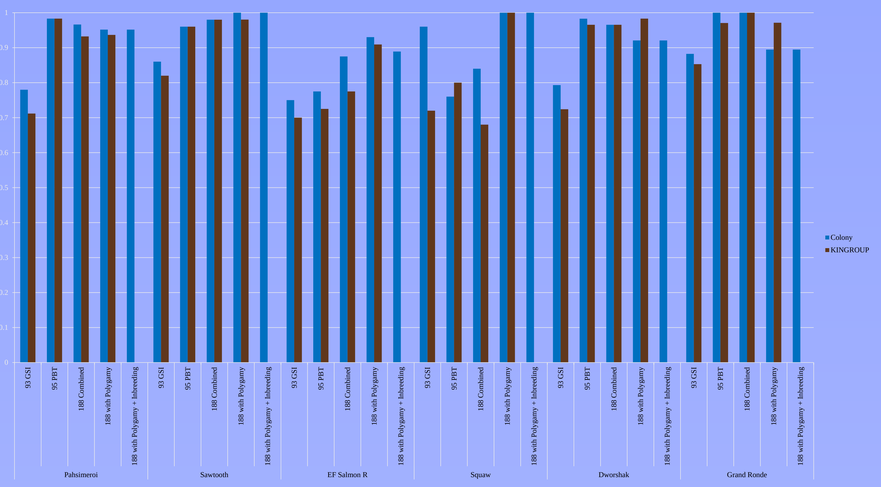

Figure 1. Accuracy of COLONY and KINGROUP at re-creating known full-sibling families. Results are grouped by stock, then by marker set and program settings used.

Methods

A novel panel of 188 steelhead SNPs were used; 93 SNPs optimized for the Columbia basin Genetic Stock Identification (GSI) program and 95 SNPs optimized for the Snake River Parentage Based Tagging (PBT) program. Six groups of 93 hatchery broodstock juveniles with known parentage were sampled for this project – Pahsimeroi, Sawtooth, East Fork Salmon River, Squaw Creek, Dworshak and Grande Ronde Lyons Ferry. Each hatchery stock of juveniles was analyzed using the 93 GSI and 95 PBT SNP marker sets separately as well as with the combined 188 marker set.

Two Programs were compared for recreating full-sibling families from offspring genotypes: COLONY 2.0 (Jones and Wang 2010) and KINGROUP (Konovalov et al. 2004). COLONY 2.0 runs were performed assuming monogamy without inbreeding, polygamy without inbreeding, and polygamy with inbreeding for the purpose of comparison. KINGROUP runs were performed using the full sibling primary hypothesis and an unrelated null hypothesis. For the KINGROUP polygamy runs, the null hypothesis was set to half siblings.

COLONY 2.0 was also compared to the program LDNE (Waples and Do 2008) for the ability to estimate N_e from offspring genotypes. COLONY N_e estimates were obtained from the run settings described previously. LDNE N_e estimates were obtained with the default program settings using the 0.02 lowest allele frequency output.

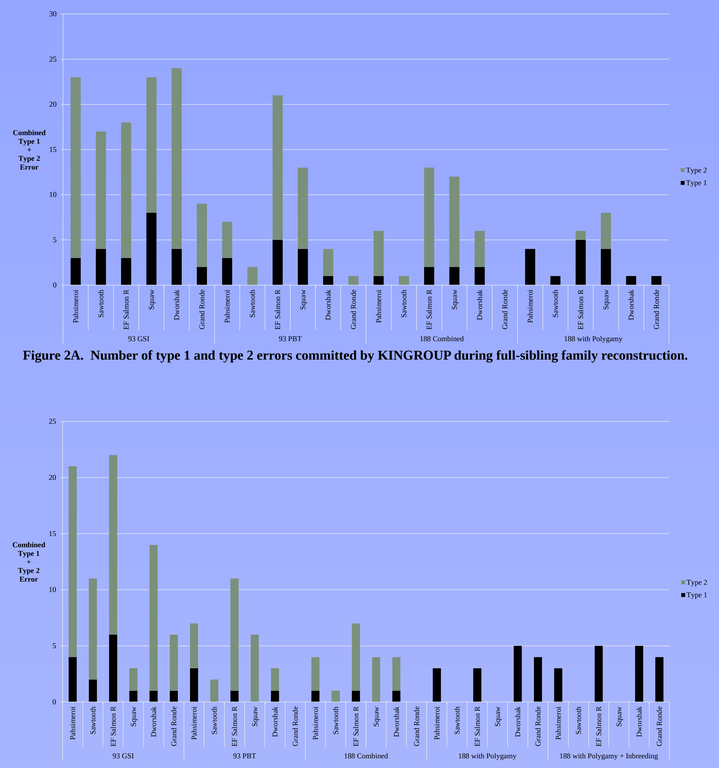

Figure 2A. Number of type 1 and type 2 errors committed by KINGROUP during full-sibling family reconstruction.

Figure 2B. Number of type 1 and type 2 errors committed by COLONY during full-sibling family reconstruction.

References

Jones, O.R. and J. Wang. 2010. COLONY: A program for parentage and sibship inference from multilocus genotype data. Molecular Ecology Resources 10:551-555.

Konovalov, D.A., C. Manning, and M.T. Henshaw. 2004. Kingroup: A program for pedigree relationship reconstruction and kin group assignments using genetic markers. Molecular Ecology Notes 4:779-782.

Laurie-Ahlberg, C.C. and B.S. Weir. 1979. Allozymic variation and linkage disequilibrium in some laboratory populations of *Drosophila melanogaster.* Genetics 92:1295-1314.

Luikart, G., N. Ryman, D.A. Tallmon, M.K. Schwartz, and F.W. Allendorf. 2010. Estimation of census and effective population sized: the increasing usefulness of DNA-based approaches. Conservation Genetics 11:355-373.

Waples, R.S. and C. Do. 2008. LDNE: A program for estimating effective population size from data on linkage disequilibrium. Molecular Ecology Resources 8:753-756.

Waples, R.S. and Waples, R.K. 2011. Inbreeding effective population size and parentage analysis without parents. Molecular Ecology Resources 11:162-171.

Acknowledgements

This work is part of a BPA funded collaborative project with the Columbia River Inter-Tribal Fish Commission's genetics lab in Hagerman, Idaho to evaluate and implement new genetic technologies for managing and conserving steelhead and Chinook salmon populations in the in the Snake River basin (2010-00-026 and 2010-00-031).

Results

Results show COLONY to be more effective than KINGROUP at recreating full-sibling families from offspring genotypes (Figure 1). As expected, the 188 combined marker set proved the most effective, followed by the 95 PBT set and the 93 GSI set. For both COLONY and KINGROUP, factoring in polygamy had a beneficial effect on the accuracy for nearly every group. Factoring inbreeding into COLONY runs had either a detrimental effect or none at all. However, only looking at accuracy measures can be misleading as type 1 and type 2 errors can cancel each other out in the final estimate of the number of full-sibling families. Figures 2A and 2B show the combined number of type 1 and type 2 errors committed by KINGROUP and COLONY runs, respectively. Type 1 error is defined as a test rejecting a true null hypothesis (not grouping two full-sibling individuals as full-siblings), while type 2 error is defined as a test accepting a false null hypothesis (grouping two individuals who are not full-siblings). Type 1 and type 2 errors are equally undesirable as an equal number of each would result in a technically 100% accurate estimate of full-sibling families.

COLONY was also more effective than LDNE at estimating N_e from offspring genotypes (Figure 3). Although COLONY underestimated N_e at times while overestimating it at others, the known N_e value always fell within COLONY's 95% confidence intervals. Factoring polygamy in to COLONY runs always resulted in an underestimate of Ne. In every case, LDNE provided an underestimate of N_e (Figure 3).

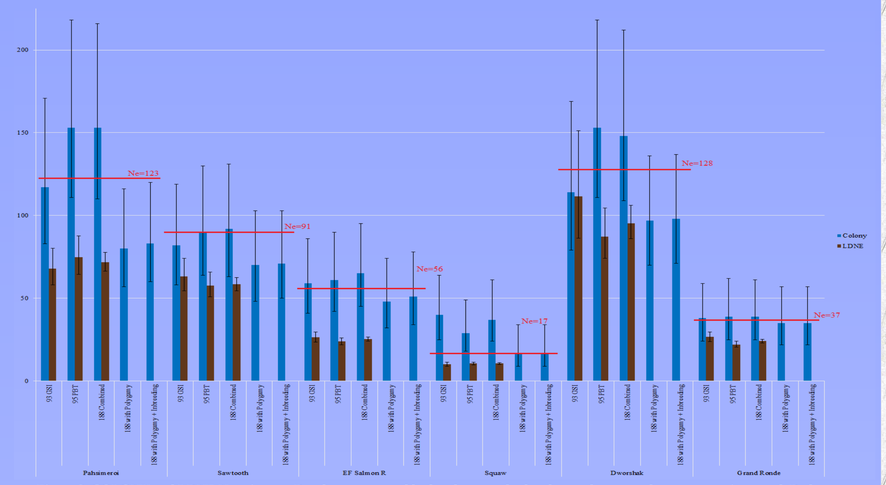

Figure 3. COLONY and LDNE N_e estimates. Results are grouped by hatchery stock, then by marker set and program settings used. Red lines indicate known N_e values for each hatchery stock. Error bars represent 95% CI reported by COLONY and LDNE.

Discussion

Inbreeding effective population size refers to the size of an ideal population that would allow the same accumulation of inbreeding as the actual population of interest. Inbreeding occurs when an offspring inherits two copies of a gene from its parents which are identical by descent (IBD)- that is, they are both directly descended from a single allele present in one of the founders of that population (perhaps the parents are cousins and each inherited the particular allele from the same grandfather). Inbreeding effective population size is thus a measure that emphasizes the effect that small population size has on the chances of relatives mating with each other. Such matings lead to a loss of heterozygosity in the population. Thus, this effective size gives an indication of the likely loss of heterozygosity across all alleles in the population.

Single-sample estimators of N_E provide an estimate of the effective number of parents that produced the progeny from which the sample is drawn, and hence, they can be associated with the inbreeding effective size (Laurie-Ahlberg & Weir 1979). Waples and Waples (2011) suggest that a single-sample estimator of contemporary inbreeding effective population size can be obtained from a random sample of progeny if sibling relationships can be accurately determined within the sample. Importantly, they also suggest that it is not necessary to know the number of adults that produced no offspring in order to provide an unbiased estimate of inbreeding effective population size. Initial results, especially those from COLONY, using these new SNP marker sets developed for steelhead are promising. They demonstrated high accuracy in estimating the number of families contributing to juvenile offspring sample sets and they provided reasonable estimates of inbreeding effective population size for the six hatchery populations tested.

Further testing under more varied conditions is necessary before it can be determined whether these methodologies can be applied to wild populations. Specifically, additional testing is needed under more complex mating system scenarios; when multiple cohorts are present; when a larger number of parents contribute offspring; and with varying sizes of sampled offspring.